DOCUMENT RESUME

ED 339 716 TM 017 542

AUTHOR Barcikowski, Robert S.; Elliott, Ronald S.
TITLE Exploratory Multivariate Analysis of Variance: Contrasts and Variables.
PUB DATE Oct 91
NOTE 27p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (12th, Chicago, IL, October 17-19, 1991).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Correlation; *Discriminant Analysis; *Mathematical Models; *Multivariate Analysis; Research Methodology; *Statistical Significance
IDENTIFIERS *Exploratory Studies; *Roy Bose Confidence Intervals

ABSTRACT

The contribution of individual variables to overall multivariate significance in a multivariate analysis of variance (MANOVA) is investigated using a combination of canonical discriminant analysis and Roy-Bose simultaneous confidence intervals. Difficulties with this procedure are discussed, and its advantages are illustrated using examples based on the following four data sets with different characteristics: (1) one-way MANOVA with three dependent variables and five subjects in each treatment; (2) 3 treatments, 4 dependent variables, and 12 subjects in each treatment; (3) one-way MANOVA with eight treatment levels and two dependent variables; and (4) 3 treatment groups and 3 variables with 125, 90, and 25 subjects, respectively. The proposed process--Roy/STP (moderate)--identifies treatment mean contrasts and dependent variables worthy of further consideration. Important parts of the process discussed can be gleaned from standard packaged statistical routines, but the step of using the Roy-Bose confidence intervals is left to the researcher. A program to complete the procedure is available in SAS and Fortran. Sixteen references are listed. An appendix illustrates the confidence interval procedure. (SLD)

Exploratory Multivariate Analysis Of Variance: Contrasts And Variables

Robert S. Barcikowski
Ronald S. Elliott

Ohio University

Paper Presented at the Annual Meeting of the Mid-Western Educational Research Association, Chicago, October, 1991.

Abstract

The contribution of individual variables to overall multivariate significance in a MANOVA is investigated using a combination of canonical discriminant analysis and Roy-Bose simultaneous confidence intervals. Difficulties with this procedure are discussed and its advantages are illustrated using examples based on four data sets with different characteristics. A call for research in this area is issued through a list of substantive problems.

Important parts of the process discussed can be gleaned from standard packaged statistical package routines e.g., BMDP, SAS, SPSSX, etc. but the step of using Roy-Bose confidence intervals is left to the researcher. A program to complete the discussed procedure, available in SAS and Fortran, will be made available to interested Mid-Western Educational Researchers.

Exploratory Multivariate Analysis Of Variance: Contrasts And Variables

Objectives

The purpose of this paper is to provide educational researchers with a means to evaluate the contribution(s) of individual variables to overall multivariate significance in exploratory multivariate analysis of variance (MANOVA). The process used (called Roy/STP(moderate)) identifies treatment mean contrasts and dependent variables worthy of further investigation. The process, however, has several areas where research is needed for further support. A second purpose of this paper is to suggest substantive areas for further research to support the use of the Roy/STP(moderate) procedure.

Perspectives

Traditionally many text book authors (e.g., Barcikowski, 1983; Cliff, 1987; Stevens, 1986) have been critical of use of Roy-Bose Simultaneous Test Procedure (STP) for considering individual variable contribution(s) to overall MANOVA significance because of the inherently conservative nature of this procedure. This is caused by the unlimited number of possible contrasts the procedure allows the researcher to consider while holding Type I error constant. Cliff (1987) put it very well by saying:

> *"it [Roy-Bose] sets the probability at alpha that any one of every conceivable contrast among groups carried out on every conceivable combination of variables will lead to rejection of the corresponding null hypothesis when all the means are equal on all the variables in the population". (p. 147)*

However, using prior knowledge of what combination of variables should be considered and using the standardized discriminant function coefficients to identify contrasts on treatment means, it will be shown that detection of the contribution of individual variables to the overall multivariate significance can be enhanced using the Roy-Bose STP.

Background

The process we illustrate is described by Bird and Hadzi-Pavlovic (1983), Hand and Taylor (1987), Harris (1985), and Lutz and Cundari (1989), but is not generally well known

(Elliott and Barcikowski, 1990) or is ignored by educational researchers (e.g., Stevens, 1986, p. 121). The process involves using canonical discriminant information to identify important variables and contrasts to be investigated using the Roy-Bose STP. As Bird and Hadzi-Pavlovic (1983) indicate, since there is multivariate significance, there is at least one contrast on the means and one on the variables which will be found to be significant. Before we illustrate how this contrast and its important variables can be found, let us consider two common misconceptions involving statistical tests in MANOVA.

Two Common Misconceptions In MANOVA

Use the Roy-Bose STP testing one dependent variable at a time. A common misconception in post hoc MANOVA analysis is that, following a significant omnibus multivariate test, the Roy-Bose STP is to be used to examine the significance of **each dependent variable alone** for treatment differences. This misconception of considering each dependent variable alone has been perpetuated by text book authors (e.g.,Barcikowski, 1983; Morrison, 1990; Stevens, 1986) who only present the Roy-Bose STP in this fashion.

The latter point is illustrated in Appendix A where Barcikowski (1983) used the Roy-Bose STP on each of two dependent variables. The value of Appendix A is that it illustrates in general how one calculates Roy-Bose simultaneous confidence intervals (SCI's).

From Appendix A you can see that the Roy-Bose STP is usually conducted for a non-repeated measures MANOVA by establishing a confidence interval about a contrast

$$\underline{\Psi} = \underline{a}'B\underline{r}. \qquad (1)$$

In equation (1), the vector $\underline{a}$ contains contrast coefficients, B is the matrix containing the population treatment (cell) means on all dependent variables, and $\underline{r}$ is the vector which allows the selection of the variables to be tested. In the next section we discuss possibilities for selecting the elements of $\underline{a}$ and $\underline{r}$, but you can see there is no reason to select elements for $\underline{r}$ that would only select one variable.

Pillai's trace is the statistic of choice to test the overall MANOVA null hypothesis. A second common misconception, based on a robustness study by Olsen (1974), is that Pillai's trace should be used to test the omnibus multivariate null hypothesis. This misconception is held because Olsen provided evidence that Pillai's trace (V) was robust to violations of MANOVA assumptions, and that Wilks'

Lambda (W), Hotelling's Trace (T), and Roy's Largest Root (R) were not as robust, especially given heterogeneous variance-covariance matrices.

We do not dispute Olsen's results. Instead, we point to results by Gabriel (1968) (discussed by Bird and Hadzi-Pavlovic (1983), and Harris (1985)) which indicate that, except for Roy's Largest Root, the multiple-root statistics are not **consonant** with their STP's. A significant omnibus test is said to be consonant with its follow-up STP when the STP can find at least one significant contrast.

We have found that while Pillai's trace performs as Olsen indicated in testing the omnibus MANOVA null hypothesis, this statistic has very low power as a follow-up STP, too often failing to detect any contrast as significant. Bird and Hadzi-Pavlovic (1983) support this contention when they indicated that:

> *The degree of conservatism associated with the V STP in unrestricted analyses is moderate in the bivariate case [two dependent variables] but so extreme elsewhere that the V STP can be rejected outright [as a choice for the STP following an omnibus MANOVA V test]. (p. 175)*

Bird and Hadzi-Pavlovic's "unrestricted" case refers to the situation where the vector **r**, discussed above, can contain any coefficients as opposed to their "moderately restricted" case where the coefficients in **r** were restricted to 1's, 0's and -1's on the following basis:

> *Variate coefficients were simplified by dividing all standardized discriminant function coefficients by the value of the largest coefficient, then replacing each rescaled coefficient by the nearest integer. (p. 173)*

Their "strongly restricted" case consisted of the **r** vector with all 0's and one 1. When Bird and Hadzi-Pavlovic examined this case the conservativeness of the Roy-Bose procedure that is described by most researchers was observed.

Roy/STP(moderate)

In what follows we define "Roy/STP(moderate)" to indicate an analysis where: (1) Roy's largest root is used to test the omnibus MANOVA null hypothesis; (2) if the latter omnibus test is significant, a Roy-Bose STP based on Roy's largest root is used with moderately restricted coefficients (as defined above) to test contrasts based on the normalized standardized discriminant function deviation means. (The procedure to find the latter contrasts is described in our

Method section.)

Bird and Hadzi-Pavlovic compared the type I error and power of the test of the omnibus MANOVA null hypothesis and of follow-up STP's (unrestricted, moderately restricted, strongly restricted) using Roy's largest root and Pillai's trace. The situations they considered were those favorable to Pillai's trace, i.e., a diffuse noncentrality structure and heterogeneous variance-covariance matrices.

Given four or six dependent variables, Bird and Hadzi-Pavlovic found that when power for the overall multivariate test was approximately .70 for the omnibus V statistic, power for the unrestricted V STP analyses was less than .20 with most power values less than .08! Power for the moderately restricted V STP analyses, which are what most researchers would probably use, was even less.

As expected from Olsen's (1974) results, given the diffuse noncentrality structure, the omnibus R statistic had less power than the omnibus V statistic, and the omnibus R statistic had higher type I error rates given heterogeneous variances. But, when the follow-up STP's with moderate restrictions were compared, the R statistic out performed the V statistic in having a less conservative type I error and greater power. The type I error rate of the Roy/STP(moderate) was close to the nominal .05 value except under the most extreme violation of variance-covariance heterogeneity when it was close to .07. The power of the Roy/STP(moderate) was reasonably close to the power of the omnibus test (about .60) but somewhat low for the STP's with four and six dependent variables (about .30).

Further Research: Substantive Problems

Based on the presentations of Bird and Hadzi-Pavlovic (1983), Gabriel (1968),and Harris (1985), we concluded that in MANOVA's:

(1) STP's using moderately restricted coefficients should become a focus of study because these coefficients are more readily interpretable. Currently, most researchers simply have an $\underline{r}$ vector containing 0's and one 1. In multivariate analyses STP's which consider several variables, e.g., contain an $\underline{r}$ vector with several 1's, a possible -1, and 0's, are reasonable.

(2) Type I error and power should be studied when Roy/STP(moderate) is the mode of analysis and the MANOVA assumptions have been met.

(3) Type I error and power should be studied when Roy/STP(moderate) is the mode of analysis and the MANOVA assumptions have not been met.

(4) The assumption violations identified in (3) should be compared with those found in practice. That is, do the violation conditions occur or are they unnessary for us to worry about (e.g., 30 to 1 variance ratios)?

(5) The limit to the number of dependent variables with which Roy/STP(moderate) will work well should be studied. Here consideration should also be given to the "all subsets" procedures described by McKay and Campbell (1982). These procedures address the problem *of assessing the relative importance of subsets of a given set of variables and, more particularly, of selecting those variables which are in some sense adequate for discrimination* (McKay and Campbell, 1982, p. 1).

(6) Researchers should begin to use Roy/STP(moderate) cautiously as a tool to help identify important contrasts and important variables in a MANOVA.

Method

The Roy/STP(moderate) analysis proceeds as follows for each source of variation in a MANOVA:

(1) Calculate the four most popular multivariate tests: Pillai's trace (V), Wilks' Lambda (W), Hotelling's Trace (T), and Roy's Largest Root (R). Stop here when Roy's largest root is not significant at your preselected level of significance (α). (Look for discrepancies among these omnibus tests which might indicate violations of the MANOVA assumptions or the need for a larger sample size (n) when a diffuse noncentrality structure is present.)

(2) Determine the significant discriminant functions (Stevens, pp. 234-235). (Although Harris (1976) argues against the legitimacy of this test of the discriminant functions, and it is based on Wilks' lambda, we use it to identify potential variables or combinations of variables to include in the Roy-Bose STP where the experimentwise error is controlled.)

(3) Normalize and standardize the coefficients of the discriminant functions and then convert them to coefficients that meet the moderate restriction definition.

(4) Calculate a set of contrast coefficients based on the discriminant function deviation means and then normalize these coefficients.

(5) For main effects only, have our Fortran program (SNOOPY.DAT) consider different possible contrasts. Contrasts for other sources of variation, e.g., interactions, would be based on the information in (4) and input by the user on a second run.

Given the contrasts that are developed at step (4), there are a variety of ways that meaningful contrast coefficients can be constructed for main effects. We begin by forming a contrast consisting of a positive one for each positive contrast coefficient and a negative one for each negative contrast coefficient. Then all possible contrasts are formed which will use these coefficients in combinations of one positive and one negative, two positives and two negatives, etc., realizing that a valid contrast must have coefficients which sum to zero. (These contrast coefficients are placed in the vector **a**′ of equation (1).)

To illustrate consider the following hypothetical vector, **a**′, found as in step (4):

-.25, -.38, -.45, .23, .11

The combination, -1 -1 -1 1 1 is formed which does not sum to zero. If the sum were zero, this vector would have been the first contrast. We then use all possible subsets of one negative and one positive, then two negatives and two positives, etc. The following list of contrasts would be formed.

-1	0	0	1	0
-1	0	0	0	1
0	-1	0	1	0
0	-1	0	0	1
0	0	-1	1	0
0	0	-1	0	1
-1	-1	0	1	1
0	-1	-1	1	1
-1	0	-1	1	1

In SNOOPY.DAT the above contrasts would be generated with Roy-Bose SCI's computed for each.

Also incorporated in the program is a contrast generated using a method formulated by Hollingsworth (1978). Using her scheme on the above vector one

would first count the negatives and positives. Then each positive component is divided by the number of positives and each negative component is divided by the number of negatives. Using the above contrast coefficients the following contrast would be tested:

$-1/3(\mu 1 + \mu 2 + \mu 3) + 1/2(\mu 4 + \mu 5)$.

This contrast has the following coefficients:

$-1/3 \quad -1/3 \quad -1/3 \quad 1/2 \quad 1/2$.

(6) Combine the variable selector vector ($\underline{r}$) found in step (3) with the contrast coefficients ($\underline{a}$) found in step (5) to determine which groups differ on what variable combinations using the Roy-Bose STP.

(7) Based on other information available to the researcher, e.g., an understanding of the relationships among the variables and of possible treatment differences, consider other possible selections of the variables and contrasts. At this stage SNOOPY.DAT may be placed in a mode where only selected $\underline{r}$ and $\underline{a}$ vectors are combined to test hypotheses of interest.

Four Examples

In this section we illustrate the above procedure using four data sets with different characteristics. These examples illustrate the potential merit of the research we advocate on the Roy/STP(moderate).

Example One: Wilkinson

Our first example was initially presented by Wilkinson (1975) to illustrate the need for a variety of multivariate procedures to analyze multivariate data. The data are for a one-way MANOVA with three dependent variables and 5 subjects in each treatment. Wilkinson indicated that the covariance structure presented by his data was not unusual for real data and he cited data presented by Jones (1966, p. 256) that had a similar structure. Stevens (1986, pp. 193-194) uses the same data in an exercise to illustrate the conservativeness of the Roy-Bose STP.

Both Wilkinson (1975, p. 411) and Stevens (1986, p. 194) provide Table 1 to illustrate the weakness of the Roy-Bose STP (Multivariate Intervals) as compared to the Scheffe STP (Univariate Intervals).

In Table 1 the Roy-Bose STP finds no variable significant for any contrast, and the Scheffe STP finds a significant difference between treatments A and B on variables 1 and 2 and between treatments B and C on variable 1. However, the Roy-Bose contrasts have been placed in their worst possible light in that they have been limited to the case of considering only one variable at a time (i.e., Bird and Hadzi-Pavlovic's most restricted case).

The output from SNOOPY.DAT contains the omnibus test results shown in Table 2. These results indicate that all four omnibus multivariate tests would reject the overall null hypothesis ($\alpha = .05$).

The omnibus univariate tests indicated that variables 1 and 2 significantly differentiated between the groups ($p < .005$ and .007, respectively) and that variable 3 did not ($p < .178$).

Table 1

Shortest 95% Confidence Intervals for Linear Contrasts Between Groups on Responses

Contrast	Variable	Multivariate Intervals	Univariate Intervals
A-B	1	$-.1 \le 2.4 \le 4.9$	$.7 \le 2.4 \le 4.1$
	2	$-.3 \le 3.4 \le 7.1$	$.9 \le 3.4 \le 5.6$
	3	$-4.4 \le -1.4 \le 1.6$	$-3.4 \le -1.4 \le .6$
A-C	1	$-1.9 \le .6 \le 3.1$	$-1.1 \le .6 \le 2.3$
	2	$-2.5 \le 1.2 \le 4.9$	$-1.3 \le 1.2 \le 3.7$
	3	$-3.8 \le -.8 \le 2.2$	$-2.8 \le -.8 \le 1.2$
B-C	1	$-4.3 \le -1.8 \le .7$	$-3.5 \le -1.8 \le -.1$
	2	$-5.9 \le -2.2 \le 1.5$	$-4.7 \le -2.2 \le .3$
	3	$-2.4 \le .6 \le 3.6$	$-1.4 \le .6 \le 2.6$

Note. Estimates of the contrasts are given at the center of the inequalities.

Table 2

The Omnibus Multivariate Test Results for the Wilkinson Data

Test Name	Value	App. F	p-Value
Pillai	.95	3.29	.018
Hotelling	2.99	4.48	.006
Wilks	.21	3.92	.009
Roy	.73	----	.011[a]

[a] Actual probability value based on Harris (1985).

Only the first discriminant function was found to be significant ($p < .009$) and its normalized and standardized coefficients were:

.44 -.79 .43.

This led to the consideration of the variable selection vector:

$\mathbf{r}' = [1.00 \quad -1.00 \quad 1.00]$.

The normalized contrast coefficients were:

-.72 .70 .02.

This led to the contrast:

$\Psi = 1\mu 1 - 1\mu 2 + 0\mu 3$

with contrast vector $\mathbf{a}' = [1 \ -1 \ 0]$.

This comparison of treatments 1 and 2 using all three variables yielded the following significant ($\alpha = .05$) Roy-Bose confidence interval:

$-4.60 \leq -2.4 \leq -0.20$,

with the estimated contrast at the center.

When we noticed that there was a high correlation between variables 1 and 2 ($r = .87$), we found that the contrast

$\Psi = 1\mu 1 - 1\mu 2 + 0\mu 3$,

with variable selector $\mathbf{r}' = [0 \ -1 \ 1]$,

was significant using the Roy-Bose STP ($\alpha = .05$). The Roy-Bose simultaneous confidence interval with the estimated contrast at the center was:

$$-8.88 \leq -4.8 \leq -0.72.$$

Care must be taken when a negative value is found in the vector $\underline{r}$. The negative value indicates that a suppressor variable is present, but in our example the vector

$$\underline{r}' = [0 \; 1 \; -1]$$

would yield the same Roy-Bose simultaneous confidence interval with the signs reversed, e.g.,

$$0.72 \leq 4.8 \leq 8.88.$$

It is left to the researcher to decide if variable 1 or variable 2 should play the role of the suppressor variable.

Our results using Roy/STP(moderate) are quit different from those found by Stevens and Wilkinson because we allowed the Roy-Bose STP to consider more than one variable. One interpretation is that variables two and three differentiate between treatments 1 and 2 with variable 3 acting as a suppressor variable, and that variable 1 can be dropped from the analysis. Wilkinson arrived at this same conclusion using a variety of analyses.

Example 2: Bird and Hadzi-Pavlovic

Bird and Hadzi-Pavlovic (1983) provide an example data set composed of three treatments, four dependent variables and twelve subjects in each treatment. In this example all of the omnibus multivariate tests were significant ($p < .005$) and the omnibus univariate tests are shown in Table 3.

Table 3

Per Variable Omnibus Univariate Tests For The Bird and Hadzi-Pavlovic Data

Variable	F	p-Value
1	5.16	.011
2	3.00	.064
3	3.36	.047
4	4.44	.020

An interesting aspect of these data is that while Pillai's trace is significant, using it in a Roy-Bose STP you will not find any contrast to be significant.

The first two discriminant functions were found to be significant ($\alpha = .05$), however Bird and Hadzi-Pavlovic only discuss the first function. The first discriminant function's normalized and standardized coefficients were:

-.074 .248 .634 .729.

This led to the consideration of the variable selection vector:

$\underline{r}' = [0.0 \quad 0.0 \quad 1.0 \quad 1.0]$.

The normalized contrast coefficients were:

-.453 .815 -.362.

This led to the contrast:

$\Psi = -.5\mu 1 + 1\mu 2 - .5\mu 3$

with contrast vector $\underline{a}' = [-.5 \quad 1 \quad -.5]$.

This comparison of the average of treatments 1 and 3 verses treatment 2 using variables 3 and 4 was found to be significant ($\alpha = .05$), as were the comparisons of treatments 1 and 2, and 2 and 3 using these same variables.

Bird and Hadzi-Pavlovic make this insightful comment concerning this analysis:

> *This example illustrates the lack of relevance of an overall MANOVA test in analyses based on univariate tests on individual measures. The first variate has the largest F ratio ($F_1 = ... = 5.16$), but it makes virtually no contribution to the first discriminant function, defined by [$\underline{r}$]. The first discriminant function is entirely responsible for the statistical significance of the overall R test and largely responsible for the significance of the overall V test. (p.172)*

Example 3: Andrews and Herzberg

Andrews and Herzberg (1985) provide data for a one-way MANOVA with eight treatment levels and two dependent variables. The four omnibus multivariate tests are significant ($p < .0001$),

as are the two omnibus univariate tests ($p < .005$).

Only the first discriminant function was found to be significant ($p < .000$) and its normalized and standardized coefficients were:

.52 .85.

This led to the consideration of the variable selection vector:

$\underline{r}' = [1.00 \quad 1.00]$.

The normalized contrast coefficients were:

.39 .33 .43 -.33 -.53 .20 -.22 -.28.

This led to the contrast:

$\Psi = 1\mu 1 + 1\mu 2 + 1\mu 3 - 1\mu 4 - 1\mu 5 + 1\mu 6 - 7\mu 1 - 8\mu 2$,

with contrast vector $\underline{a}' = [1 \quad 1 \quad 1 \quad -1 \quad -1 \quad 1 \quad -1 \quad -1]$.

This comparison of treatments 1, 2, 3, and 6 with treatments 4, 5, 7, and 8 using both variables yielded the following significant ($\alpha = .05$) Roy-Bose confidence interval:

$-71.62 \leq -49.39 \leq -27.16$,

with the estimated contrast at the center.

In taking these contrast coefficients in pairs of three's (i.e., three positive and three negative coefficients), two's and one's, plus the Hollingsworth coefficients, SNOOPY.DAT considered 69 contrasts, 64 of which were significant. When the first variable was taken alone no significant contrasts were found. When the second variable was taken alone 60 significant contrasts were found.

When both dependent variables were considered the following contrasts were found to be significant but were not significant when only the second variable was considered alone:

1	0	0	0	0	0	0	-1
0	1	0	0	0	0	0	-1
0	0	1	0	0	0	0	-1
0	1	0	0	0	1	-1	-1
1	0	0	0	0	0	-1	0
0	0	1	0	0	0	-1	0.

When the second variable was taken alone the following contrasts were found to be significant but were not significant when both variables were considered together.

```
0  0  0 -1  0  1  0  0
0  1  0 -1  0  0  0  0
```

These data illustrate that the SNOOPY.DAT output can provide researchers with a wide variety of choices for interpretation.

Exampie 4: Hand and Taylor

Hand and Taylor (1987, pp. 97-116) provide data from a study on alcohol relapse. The data consist of three treatment groups and three variables. The treatment groups are described in Table 4. The variables, measures of vulnerability to relapse, are as follows:

> *(UM) unpleasant mood states; for example, depression;*
>
> *(ES) euphoric states and related situations; for example, celebrations and parties; and*
>
> *(LV) an area designated as lessened vigilance; for instance, a temptation to believe that one or two drinks would cause no problem. (p. 104)*

Treatment 1 had 125 subjects, treatment 2 had 90 subjects and treatment 3 had 25 subjects.

Table 4

Treatment Descriptions in the Hand and Taylor Study[a]

Treatment	Label	Description
1	long-standing	those who had a longer history of relapse of four or more times
2	recent	those who claimed to have relapsed, but no more than two or three time before
3	new	those never having previously experienced relapse after trying to give up heavy drinking

[a]Taken from p. 98 of Hand and Taylor (1987).

Only the first discriminant function was found to be significant ($p < .006$) and its normalized and standardized coefficients were:

-.15 .16 -.98.

This led to the consideration of the variable selection vector:

$\mathbf{r}' = [0.0 \quad 0.0 \quad -1.00]$.

The normalized contrast coefficients were:

-.73 .05 .68.

This led to the contrast:

$\Psi = -1\mu 1 + 0\mu 2 + 1\mu 3$

with contrast vector $\mathbf{a}' = [1 \ 0 \ -1]$.

This comparison of treatments 1 and 3 using the third variable yielded the following significant ($\alpha = .05$) Roy-Bose confidence interval:

$-3.57 \leq -1.92 \leq -0.27$,

with the estimated contrast at the center.

The Hollingsworth contrast [1 -.5 -.5], testing only the third variable, was also significant with Roy-Bose confidence interval ($\alpha = .05$):

$-2.63 \leq -1.47 \leq -0.31$.

Hand and Taylor's results were based on *a priori* contrasts, however, the Hollingsworth contrast yielded results that were consistent with their's. That is, they concluded (p. 106) ...*that there is a difference between the mean vector of group 1 and the mean vector of groups 2 and 3.* They also indicated that (p. 108): ... *the primary weighting is on the third response, LV, and we conclude that this measure alone will give almost as good a discrimination.*

This example illustrates a condition under which the Roy-Bose STP can be potent (powerful) using one dependent variable, that is, when the normalized and standardized discriminant function coefficients indicate that a single variable is responsible for treatment differences.

Conclusions

The process offered in this paper gives the researcher an extra tool in the ongoing development of methods to analyze the contribution of individual variables to overall MANOVA significance. We feel that Roy/STP(moderate), or a modification of it, has great potential and we hope that methodologists will take up our challenge to carry out research on this important topic.

Warnings

We feel that we can not leave this topic without considering a list warnings to perspective users of Roy/STP(moderate).

(1) Because of the difficulties inherent in finding stable discriminant function coefficients, Roy/STP(moderate) may not work well with small sample sizes and/or large numbers of dependent variables. However the work of Hadzi-Pavlovic (1983) indicates that power can be studied, and our own observation is that researchers frequently do not use large numbers of dependent variables in studies involving MANOVA's. Also, note some of the small cell sample sizes in some of our examples where Roy/STP(moderate) worked quite well.

(2) Stevens (1986, p. 121) indicates that the variables provided by a discriminant analysis may prove to be uninterpretable to the researcher. We find this difficult to believe, especially given moderate restrictions on the coefficients, because *post hoc* most researchers can find an interpretation for just about anything.

Educational Importance of the Study

The need for a method to consider contributions of individual variables to an overall multivariate significant result is evident in the literature (Elliott and Barcikowski, 1990). Current writing on the subject ranges from the extreme that there is no true multivariate analysis to actually contributing causes of the overall multivariate significance to individual variables using univariate results. Other efforts, including this writing, lie between these two results.

References

Andrews, D. F. & Herzberg, A. M. (1985). *Data: A collection of problems from many fields for the student and research worker*. Springer-Verlag: New York.

Barcikowski, R. S.(ED.)(1983). *Computer package and research design, with annotations of input and output from the BMDP, SAS, SPSS, and SPSSX statistical packages. (Vol.1 BMDP, VOL. 2:SAS, Vol. 3:SPSS and SPSSX)*. Lanham,MD: University Press of America.

Bird, K. D. & Hadzi-Pavlovic, D.(1983). Simultaneous test procedures and the choice of a test statistic in MANOVA. *Psychological Bulletin, 93(1)*, 167-178.

Cliff, N.(1987). *Analyzing multivariate data*. Orlando, FL: Harcourt Brace Jovanovich.

Elliott, R. & Barcikowski, R.(1990, October), *Causes of multivariate significance in a multivariate analysis of variance*. Paper presented at the meeting of the Midwestern Educational Research Association, Chicago, IL.

Jones, L. V. (1966). Analysis of variance in its multivariate developments. (in R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology*. Chicago: Rand McNally.

Hand, D. J. & Taylor, C.C. (1987). Multivariate analysis of variance and repeated measures. Chapman & Hall: New York.

Harris, R. J.(1976). The invalidity of partitioned U tests in canonical correlation and multivariate analysis of variance. *Multivariate Behavioral Research, 11*, 353-365.

Harris, R. J.(1985). *A primer of multivariate statistics*. (2nd ed.). Orlando, FL: Academic Press, Inc.

Hollingsworth, H. (1978). *The coefficients of a maximum contrast as interpretable statistics*. A paper presented before the annual meeting of The American Educational Research Association, Toronto, Ontario, March.

Lutz, J. G. & Cundari, L. A. (1989). Determining the most significant parametric function for a given linear multivariate hypothesis. *Journal of Educational Statistics, 14(1)*, 21-28.

McKay, R. J. & Campbell, N. A. (1982). Variable selection techniques in discriminant analysis i. Description. *British Journal of Mathematical and Statistical Psychology, 35,* 1-29.

Morrison, D. F. (1990). *Multivariate statistical methods.* (3rd ed.). New York, NY: McGraw-Hill.

Olson, C. L. (1974).Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association, 69,* 894-908.

Stevens, J. P.(1986). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Wilkinson, L.(1975). Response variable hypotheses in the multivariate analysis of variance. *Psychological Bulletin, 82,* 408-412.

Appendix A
An Illustration Of Roy-Bose Simultaneous Confidence Intervals

This appendix was taken from material in Barcikowski (1983). The data are shown in Figure 3.2, on the next page. In the analysis no interaction was found, but a significant difference was found between the test types ($p < .0094$). In what is shown here, the investigators are using the Roy-Bose simultaneaous confidence intervals to see if: (1) there is a significant difference on the unweighted attitude means between students who had easy and difficult tests, (2) there is a significant difference on the unweighted achievement means between students who had easy and difficult tests.

The Roy-Bose simultaneous confidence interval procedure is generally used after the multivariate null hypothesis has been rejected; it can be calculated using the following steps.

Step 1

State the general linear model for the design. The general linear model in full rank may be written as $Y = XB + E$ where,

Factor A Test Type	Factor B Instruction Method: Auditory		Heuristic		
	attitude	achievement	attitude	achievement	
Easy	6	10	5	13	weighted means
	4	7	7	11	6.00 11.67
			8	16	
			6	13	unweighted means
	M = 5.00	8.50	6.50	13.25	5.75 10.88
	attitude	achievement	attitude	achievement	
Difficult	3	6	6	7	weighted means
	3	3	5	11	3.58 5.08
	2	1	4	9	
	4	8	5.00	9.00	unweighted means
	6	5			4.06 6.39
	5	4			
	2	2			
	2	2			
	1	3			
	M = 3.11	3.78			
weighted means	3.45	4.64	5.86	11.43	
unweighted means	4.06	6.14	5.75	11.12	

Figure 3.2 *Data and means for a two factor nonorthogonal fixed-effects experimental design, where the dependent variables are student attitude towards and achievement in arithmetic.*

N is the total number of units of analysis (e.g., subjects) in the design,
P is the number of dependent variables,
L is the number of cells in the design,
Y is the $N \times P$ matrix of observed scores on the dependent variables (across all cells in the design),
X is the $N \times L$ design matrix which is composed of ones and zeros,
B is the $L \times P$ matrix of population cell means, and
E is the $N \times P$ matrix of errors (deviations of each score from its cell population mean).

Example. Our two-way multivariate general linear model may be written as $Y = XB + E$, as shown in Figure 3.3.

Step 2

State the multivariate null hypothesis. The general form of the multivariate null hypothesis is: $ABC' = D$ where,

G is the number of contrasts among the cell means that one decides to investigate, G is less than or equal to the between degrees of freedom from the source of variation under investigation.
L is the number of cells in the design,
P is the number of dependent variables,
M is the number of contrasts among the dependent variables that one decides to investigate, $M \leq P$,
A is a $G \times L$ matrix of contrasts among treatment cell means,
B is a $L \times P$ matrix of unknown population cell means on the dependent variables,
C' is a $P \times M$ matrix of contrasts among the dependent variables, and
D is a $G \times M$ matrix of known population values.

Example. In the problem we have been considering we may state the multivariate null hypothesis of no test type effects as $AB = D$, as follows:

$$
\begin{array}{ccccccc}
Y & = & X & & B & + & E \\
\begin{bmatrix}
6 & 10 \\ 4 & 7 \\ 5 & 13 \\ 7 & 11 \\ 8 & 16 \\ 6 & 13 \\ 3 & 6 \\ 3 & 3 \\ 2 & 1 \\ 4 & 8 \\ 6 & 5 \\ 5 & 4 \\ 2 & 2 \\ 2 & 2 \\ 1 & 3 \\ 6 & 7 \\ 5 & 11 \\ 4 & 9
\end{bmatrix}
& = &
\begin{bmatrix}
1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1
\end{bmatrix}
& &
\begin{bmatrix}
\mu(111) & \mu(211) \\ \mu(112) & \mu(212) \\ \mu(121) & \mu(221) \\ \mu(122) & \mu(222)
\end{bmatrix}
& + &
\begin{bmatrix}
e(1111) & e(2111) \\ e(1112) & e(2112) \\ e(1121) & e(2121) \\ e(1122) & e(2122) \\ e(1123) & e(2123) \\ e(1124) & e(2124) \\ e(1211) & e(2211) \\ e(1212) & e(2212) \\ e(1213) & e(2213) \\ e(1214) & e(2214) \\ e(1215) & e(2215) \\ e(1216) & e(2216) \\ e(1217) & e(2217) \\ e(1218) & e(2218) \\ e(1219) & e(2219) \\ e(1221) & e(2221) \\ e(1222) & e(2222) \\ e(1223) & e(2223)
\end{bmatrix}
\end{array}
$$

FIGURE 3.3 *The full rank general linear model, written as $Y = XB + E$, for the test type by instructional method experiment. Here, $\mu(pij)$ and $e(pijk)$ are subscripted with; variable p, cell i,j, and within cell k.*

$$H_{01}: \quad [1/2 \;\; 1/2 \;\; -1/2 \;\; -1/2] \begin{bmatrix} \mu(111) & \mu(211) \\ \mu(112) & \mu(212) \\ \mu(121) & \mu(221) \\ \mu(122) & \mu(222) \end{bmatrix} = [0 \;\; 0]$$

$$A \qquad B \qquad = \qquad D$$

where C' (from the general form of the null hypothesis) is the identity matrix and is not needed in the hypothesis equation.

Step 3

Let $\underline{a}'$ represent a row of A for a contrast of interest. Calculate the constant b as:

$$b = \underline{a}'(X'X)^{-1}\underline{a}.$$

Example. Let N_{ij} represent the number of subjects in cell i,j then, for our example problem we have:

$$(X'X)^{-1} = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1/9 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}$$

Since there is only one row of contrasts in A, for the hypothesis of no test type effect, we have that $A = a'$, and

$$A(X'X)^{-1}A' = 1/4 \;\; [1/2 \;\; 1/4 \;\; 1/9 \;\; 1/3] = .29861$$

so that, $b = .29861$.

Step 4

The Roy-Bose simultaneous confidence interval for the function $\underline{a}'Bc'\underline{r}$ may be written as:

where, $\underline{r}$ is an $M \times 1$ non null vector of real elements, and $H(\alpha)$ is the 100 (α) percentage point that is read from charts compiled by Heck (1960) (see Morrison, 1976, pp. 379-386). The parameters needed to find $H(\alpha)$ are:

$s = \min(G, p)$,
$m = (G - p - 1)/2$,
$n = (DFE - p - 1)/2$,
DFE = degrees of freedom for error, i.e., $DFE = N - Q - 1$, where Q is the total of the degrees of freedom for all factorial sources of variation in the design, and E is the error sums of squares and cross products matrix, E-$SSCP$. (Note that we also used E to represent the error matrix in the general model.)

Example. In our design we may desire to establish Roy-Bose simultaneous confidence intervals for the unweighted mean attitude and achievement differences found between the test type levels. To do this we must fill in the needed matrices and vectors in Step 4. Since C' is an identity matrix and may be dropped, we have that the simultaneous confidence interval becomes:

$$\underline{a}'(X'X)^{-1}X'YC'\underline{r} - \sqrt{\frac{b\,H(\alpha)}{1 - H(\alpha)}\,\underline{r}'E\underline{r}} \;\le\; \underline{a}'BC'\underline{r} \;\le\; \underline{a}'(X'X)^{-1}X'YC'\underline{r} + \sqrt{\frac{b\,H(\alpha)}{1 - H(\alpha)}\,\underline{r}'E\underline{r}}$$

$$\underline{a}'(X'X)^{-1}X'Y\underline{r} - \sqrt{\frac{b\,H(\alpha)}{1-H(\alpha)}\,\underline{r}'E\,\underline{r}}$$

$$\leq \quad \underline{a}'B\,\underline{r} \leq$$

$$\underline{a}'(X'X)^{-1}X'Y\,\underline{r} + \sqrt{\frac{b\,H(\alpha)}{1-H(\alpha)}\,\underline{r}'E\,\underline{r}}$$

Here we see that $\underline{a}'B$ (in the center of this inequality) is a row vector containing the elements:

$[\mu(111)+\mu(112)-\mu(121-\mu(122)]/2$

and

$[\mu(211)+\mu(212)-\mu(221-\mu(222)]/2.$

Therefore, if we want a simultaneous confidence interval on the unweighted attitude means we select

$$\underline{r} = \begin{bmatrix}1\\0\end{bmatrix}.$$

Now, $(X'X)^{-1}X'Y$ yields the $L \times P$ matrix of our sample cell means, i.e.,

$$(X'X)^{-1}X'Y = \begin{bmatrix}5.00 & 8.50\\6.50 & 13.25\\3.11 & 2.78\\5.00 & 9.00\end{bmatrix}$$

So that $\underline{a}'(X'X)^{-1}X'Y$ yields the vector of unweighted cell mean differences, i.e.,

$\underline{a}'(X'X)^{-1}X'Y = [1.695\ 4.985]$.

$H(.05)$ with $s=1$, $m=0$, $n=5.5$ is not available from the Heck charts, but may be found using (Morrison, p. 178)

$$H(\alpha) = \frac{(m+1)\,F(\alpha)}{(n+1)+(m+1)\,F(\alpha)}$$

where $F(\alpha)$ is Fisher's F (the $(1-\alpha)$ percentile) with $2m+2$ and $2n+2$ degrees of freedom. In our case $F(.05;2,13)=3.81$ and $H(\alpha)=3.81/(6.5+3.81)=.37$. The value of E, is (from Table 3.4)

$$E = \begin{bmatrix}29.89 & 20.72\\20.72 & 64.81\end{bmatrix}$$

Then, setting $\underline{r} = \begin{bmatrix}1\\0\end{bmatrix}$,

and substituting the preceding values into the Roy-Bose simultaneous confidence interval, we have:

$$1.695 - \sqrt{((.29861 \times .37)/(1-.37))\,29.89}$$

$$\leq \mu(11.) - \mu(12.) \leq$$

$$1.695 + \sqrt{((.29861 \times .37)/(1-.37))\,29.89}$$

where $\mu(pi.)$ is the unweighted mean of factor A, for variable p and treatment level i; the dot notation indicates that we summed across the j levels of Factor B. The final calculations yield:

$$-.59 \leq \mu(11.) - \mu(12.) \leq 3.97$$

The preceding interval contains zero and therefore no significant difference is indicated between the attitudes of those students who had different test types.

Setting $\underline{r} = \begin{bmatrix}0\\1\end{bmatrix}$,

we have the following simultaneous confidence interval for the unweighted achievement mean

differences for factor A:

$$1.13 \le \mu(21.) - \mu(22.) \le 7.87.$$

This interval does not contain zero and therefore a significant difference was found between the achievement of those students who had easy tests and those students who had difficult tests.